DOCUMENT RESUME

ED 161 953 TM 008 003

AUTHOR Maier, Milton H.; Hirshfeld, Stephen F.

TITLE Criterion-Referenced Job Proficiency Testing: A Large Scale Application. Research Report 1193.

INSTITUTION Army Research Inst. for the Behavioral and Social Sciences, Arlington, Va.
SPONS AGENCY Army Training and Doctrine Command, Fort Monroe, Va.; Office of the Deputy Chief of Staff for Personnel (Army), Washington, D.C.
PUB DATE Feb 78
NOTE 30p.

EDRS PRICE MF-$0.83 HC-$2.06 Plus Postage.
DESCRIPTORS *Criterion Referenced Tests; *Enlisted Personnel; Job Analysis; *Job Skills; Mastery Tests; *Occupational Tests; *Performance Tests; Personnel Evaluation; Test Construction; Test Interpretation
IDENTIFIERS *Army; *Skill Qualification Tests

ABSTRACT

The Army Skill Qualification Tests (SQT's) were designed to determine levels of competence in performance of the tasks crucial to an enlisted soldier's occupational specialty. SQT's are performance-based, criterion-referenced measures which offer two advantages over traditional proficiency and achievement testing programs: test content can be made public in advance of administration, and personnel management decisions involving promotion, selection, and advanced schooling can be based on performance standards instead of personnel quotas. SQT's are based on job descriptions provided in the Soldier's Manual and consist of scorable units for job tasks. A number of test items make up a scorable unit through which a soldier must demonstrate acceptable mastery of the task in question. This booklet discusses standards for SQT's, test development, and assumptions for using SQT scores. (Author/JAC)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

- ED161953

TM

Research Report 1193

ADA055664

CRITERION-REFERENCED JOB PROFICIENCY TESTING: A LARGE SCALE APPLICATION

Milton H. Maier and Stephen F. Hirshfeld

INDIVIDUAL TRAINING AND SKILL EVALUATION TECHNICAL AREA

U. S. Army

Research Institute for the Behavioral and Social Sciences

February 1978

Approved for public release; distribution unlimited.

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

TM008 003

U. S. ARMY RESEARCH INSTITUTE
FOR THE BEHAVIORAL AND SOCIAL SCIENCES

A Field Operating Agency under the Jurisdiction of the
Deputy Chief of Staff for Personnel

J. E. UHLANER
Technical Director

W. C. MAUS
COL, GS
Commander

NOTICES

DISTRIBUTION: Primary distribution of this report has been made by ARI. Please address correspondence concerning distribution of reports to: U. S. Army Research Institute for the Behavioral and Social Sciences, ATTN: PERI-P, 5001 Eisenhower Avenue, Alexandria, Virginia 22333.

FINAL DISPOSITION: This report may be destroyed when it is no longer needed. Please do not return it to the U. S. Army Research Institute for the Behavioral and Social Sciences.

NOTE: The findings in this report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER Research Report 1193	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) CRITERION-REFERENCED JOB PROFICIENCY TESTING: A LARGE SCALE APPLICATION		5. TYPE OF REPORT & PERIOD COVERED 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Milton H. Maier and Stephen F. Hirshfeld		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS U.S. Army Research Institute for the Behavioral and Social Sciences 5001 Eisenhower Avenue, Alexandria, VA 22333		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 2Q762722A764
11. CONTROLLING OFFICE NAME AND ADDRESS Deputy Chief of Staff for Personnel, Washington DC; Army Training and Doctrine Command, Fort Monroe, VA		12. REPORT DATE February 1978 13. NUMBER OF PAGES 20
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Performance-based testing
Criterion-referenced testing
Skill Qualification Tests (SQT)
Military Occupational Specialty (MOS) tests
Job proficiency testing

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The Army's Skill Qualification Test (SQT) is designed to test levels of competence in performance of the jobs crucial to an enlisted soldier's Military Occupational Specialty (MOS). Each SQT is based on task analysis of the job as described in the Soldier's Manual for that job, thus providing open knowledge of the requirements. An SQT consists of scorable units for the tasks involved in the job. A number of test items, each one covering a critical element of job performance, make up a scorable

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Block No. 20 (cont'd)

unit, which is scored GO if the soldier has demonstrated acceptable mastery of the task in question. The nature of the task determines whether both hands-on performance tests and written-performance based tests are used in the scorable units. Tests must be fair to all soldiers and feasible for standardized worldwide administration.

Test scores indicate an individual's level of competence and hence pay-grade or need for further training. Test scores also provide personnel management with specific performance data on which to base personnel decisions.

Research Report 1193

CRITERION-REFERENCED JOB PROFICIENCY TESTING: A LARGE SCALE APPLICATION

Milton H. Maier and Stephen F. Hirshfeld

Submitted by:

Milton S. Katz, Chief

INDIVIDUAL TRAINING AND SKILL EVALUATION TECHNICAL AREA

E. Ralph Dusek, Director
INDIVIDUAL TRAINING AND PERFORMANCE
RESEARCH LABORATORY

Approved By:
J. E. Uhlaner
TECHNICAL DIRECTOR

U.S. ARMY RESEARCH INSTITUTE FOR THE BEHAVIORAL AND SOCIAL SCIENCES
5001 Eisenhower Avenue, Alexandria, Virginia 22333

Office, Deputy Chief of Staff for Personnel
Department of the Army

February 1978

Army Project Number
2Q762722A764

Performance-Based Skill Evaluation

Approved for public release; distribution unlimited.

ARI Research Reports and Technical Papers are intended for sponsors of R&D tasks and other research and military agencies. Any findings ready for implementation at the time of publication are presented in the latter part of the Brief. Upon completion of a major phase of the task, formal recommendations for official action normally are conveyed to appropriate military agencies by briefing or Disposition Form.

FOREWORD

Improvement in the efficiency and economy of individual enlisted training, evaluation, and utilization is essential to maintain maximum combat readiness of the Army, and is a major concern of the Individual Training & Skill Evaluation Technical Area of the Army Research Institute for the Behavioral and Social Sciences (ARI). The present Army policy emphasizes performance-based training and testing; ARI research has made possible the development, validation, and application of performance-based, criterion-referenced Skill Qualification Tests (SQTs) as well as self-contained procedures by which Army/Test Development Agencies can construct and validate the SQTs.

The present report discusses the SQT program, its principles of test construction, and the benefits expected in its utilization. Research was accomplished under Army Project 2Q762722A764, and is directly responsive to the requirements of the Individual Training and Evaluation Directorate (ITED) of the Army Training Support Center, Fort Eustis, Virginia.

J. E. UHLANER,
Technical Director

CRITERION-REFERENCED JOB PROFICIENCY TESTING: A LARGE SCALE APPLICATION

BRIEF

Requirement:

Army training and personnel management requires job performance tests that are fair to all soldiers, feasible for worldwide administration, and measure performance on critical job tasks.

Procedure:

Procedures for developing Skill Qualification Tests (SQTs) were prepared and tried out by Army test development agencies. The procedures cover assuring that the tests have content validity and verifying that the tests are accurate measures of performance.

Results:

- Procedures for developing criterion-referenced, performance-based evaluations of task performance.
- Procedures for determining accuracy of the tests as measures of performance.
- Guidelines and self-instructional materials for developing SQTs. The procedures are designed to assure that the tests are based on realistic job requirements and that the scores reflect successful task performance (that is, they are criterion referenced). The general test content, therefore, can be open knowledge, and subsequent management decisionmaking can be based on how well soldiers attain performance standards.

Utilization:

The procedures for constructing and validating Skill Qualification Tests are in use for developing more than 1000 tests for evaluating job proficiency in the Army enlisted force. The guidelines and self-instructional materials are used to train personnel at the more than thirty Test Development Agencies on how to develop SQTs.

CRITERION-REFERENCED JOB PROFICIENCY TESTING: A LARGE SCALE APPLICATION

CONTENTS

	Page
OVERVIEW	1
BACKGROUND	1
REQUIREMENTS FOR SKILL QUALIFICATION TESTS	2
Performance Information for Training and Personnel Management	3
Job Relevance	4
DEVELOPMENT OF SKILL QUALIFICATION TESTS	4
Fairness and Feasibility of the Tests	4
Establishing a Correspondence Between Test Content and Job Tasks	7
Validating Tests Prior to Administration	8
ASSUMPTIONS FOR USING SQT SCORES	9
Help Determine Training Requirements	9
Help Select Soldiers in a Single Specialty	10
Help Select Soldiers in Merged Specialties	10
BENEFITS FROM USING CRITIERION-REFERENCED SQTs	11
Public Nature of Test Content	11
Job Performance Standards vs. Personnel Quotas	13
CONCLUSIONS	17
DISTRIBUTION LIST	19

CRITERION-REFERENCED JOB PROFICIENCY TESTING: A LARGE SCALE APPLICATION

OVERVIEW

Skill Qualification Tests (SQT) have been developed to replace Military Occupational Specialty (MOS) proficiency tests as measures of ability to perform Army enlisted jobs. SQTs are performance-based, criterion-referenced measures of job proficiency, consisting of precisely defined tests of tasks, all of which are critical and necessary to performance of the job. The criterion-referenced approach provides an explicit relationship between job requirements and test content in that job requirements dictate content of SQTs. The SQT development process requires that tests be reviewed by subject matter experts and validated on representative job incumbents to assure that test content is job relevant. Test standards of acceptable levels of performance are also based on job requirements and test content. Performance standards are based on behaviorally derived absolute scoring standards, and are not based on performance relative to other soldiers who take the test. For these reasons SQTs are justifiably viewed as criterion-referenced tests of job proficiency.

A criterion-referenced testing system offers two significant advantages not available in traditional testing programs. One is that test content can be made public in advance of administration. There are no reasons to keep test content secret in a testing program based on explicit linkages between test content and job requirements. Advance knowledge of test content results in an equitable and open system. Everyone has an equal opportunity to acquire proficiency on the specific job tasks known to be included in the test.

The second is that a criterion-referenced approach allows personnel management decisions such as promotion, selection, and advanced schooling to be based on performance standards instead of personnel quotas. In more complicated situations involving the merging or splitting of job specialties at higher skill levels, soldiers from different specialties can be compared on the basis of their levels of competence instead of their relative standing in the testing group. Criterion-referenced testing of job proficiency has opened new opportunities for both training and personnel management.

BACKGROUND

The Army has been using tests to measure job proficiency for over 15 years. These tests, called Military Occupational Specialty (MOS) proficiency tests, were designed primarily to help personnel managers in making decisions of vital importance to individuals' careers, such as proficiency pay, promotion, and assignments. The MOS tests were traditional achievement tests, consisting of 125 multiple choice items, each with four alternatives. The test content was related generally to the

domain of job performance, but there was no definitive logical correspondence between test items and specific job requirements. Each item was scored pass-fail; the total score was the number of items correct, and the total score was then used to rank order persons in each job specialty. Therefore, any referencing of test score to test content was immediately abandoned.

While such proficiency tests have use in personnel management decisions, they did not fully serve the Army training needs. Because of content limitations, lack of content-score correspondence, minimal diagnostic utility, and the long delay in providing feedback to the field (up to one year after testing), Army trainers did not find MOS tests particularly useful for determining training requirements, measuring individual and unit performance, and defining training readiness.

Army training during this same period, especially in the late 1960's and early 1970's, was undergoing a major revolution. Performance-based training and testing, based on critical job tasks and criterion-referenced standards of performance, were being implemented in entry-level training courses. Training objectives were operationally defined by the performance tests given during the course, and tests were made public to students as well as instructors. The content of these tests was always directly relevant to the job. The tests themselves were used to drive the direction of training.

Tests, because of their function in maintaining accountability, are effective instruments in bringing about institutional change. Test content helps implement doctrine about the way jobs are to be performed, and is helpful in defining training requirements and standards. The public nature of the tests helps focus attention on the critical elements of the job, enables effective use of soldiers' time in preparing for tests, and thus improves individual readiness.

So impressive was the success of performance-based training and testing that the Army made the policy decision to change from the existing mode of job proficiency testing, typically referred to as "norm-referenced, paper-and-pencil testing," to the criterion-referenced mode of proficiency testing. These new criterion-referenced tests, called Skill Qualification Tests (SQT), are having a profound impact on the entire Army community. The new testing procedures are forcing training managers, personnel managers, and research support personnel to rethink and often redefine their functions.

REQUIREMENTS FOR SKILL QUALIFICATIONS TESTS

The basic requirement of SQTs is that the tests are job relevant. The test content must be based on job requirements, and the test scores must be accurate measures of ability to perform critical job tasks.

The job relevance of SQTs is assured by basing them on Soldier's Manuals. Soldier's Manuals contain the critical job tasks, the behaviors required to perform the tasks, the job conditions, and the standards of performance. Soldier's Manuals define the jobs in that they list all the tasks soldiers in a job specialty are responsible for performing. Since SQTs are based on Soldier's Manuals, the SQTs are job relevant.

PERFORMANCE INFORMATION FOR TRAINING AND PERSONNEL MANAGEMENT

SQTs are used by both training and personnel management to help make important decisions affecting the career development of soldiers. Both training and personnel management need timely and accurate information about how well individuals are performing; training management to determine training requirements of individuals, and personnel management to help determine who to promote, reclassify, or reassign. Although both training and personnel management have a need for the same kind of information, their immediate requirements are not identical.

Training managers base their immediate training requirements on the specific tasks performed in their units. The job relevance of tests for specific assignments, therefore, is the primary consideration from this point of view and it is defined in terms of the tasks that soldiers perform in their assignments. The set of tasks performed in an assignment is generally a subset of the tasks required in a specialty. The task is a convenient unit for determining training requirements because tasks are observable, have initiating and terminating cues, and have standards of performance that can be reasonably well specified. Decisions about proficiency can be made at the task level, and training managers can identify the specific tasks on which soldiers need training. If the test measures performance on the specific tasks for which the training managers have responsibility, then the tests are serving their basic purpose.

Personnel managers are also concerned with the job performance of individual soldiers; but rather than focusing on soldiers' specific assignments, personnel managers need to know how well soldiers can perform all the tasks in a specialty. For example, performance in a specialty, such as Infantryman or Wheeled Vehicle Mechanic, cannot necessarily be inferred from the set of tasks found in any one assignment. Personnel managers, therefore, have a need for information based on a standard set of tasks for each specialty. All soldiers in a specialty need to be evaluated on the same set of tasks to enable fair decisions about which soldiers to promote, retain, or reclassify. The need for a standard set of tasks in each specialty imposes additional testing requirements for feasibility and acceptability. The test scores should not be affected by when or where the test is taken, nor by whom it is administered and scored. The testing conditions, as well as performance standards, should be standardized.

JOB RELEVANCE

The requirement for Army-wide standardization at the present state of the art in testing means that initially most of the test content is in the paper-and-pencil mode rather than hands-on performance tests. Paper and pencil tests generally lack the apparent job relevance of hands-on performance tests, and therefore an additional requirement is imposed to assure that the tests are acceptable to examinees, supervisors and commanders as valid measures of job proficiency.

Job relevance of the tests is the basic requirement for both training and personnel management, even though the definition of job relevance may have somewhat different meanings for the two purposes. For training purposes the focus is on the subset of tasks performed in the specific job assignment, whereas for personnel purposes the interest is on the entire set of tasks in the specialty.

The SQTs are designed to serve the requirements of training and personnel management. Because of their somewhat divergent immediate needs, critical issues arise in how SQTs are developed, scored, and used. These issues--notably the public nature of test content and personnel quotas as performance standards--are treated in this paper.

The next section describes the development of Skill Qualification Tests and expands on the technical requirements, managerial requirements, and practical constraints described in this section. The subsequent sections describe assumptions in scoring SQT and benefits resulting from adopting a criterion-referenced approach to SQT development. The magnitude of these benefits far outweighs the costs of developing and implementing such a large-scale program.

DEVELOPMENT OF SKILL QUALIFICATION TESTS

The Skill Qualification Testing (SQT) program is a large scale attempt to provide valid and efficient measures of job proficiency. This section describes the process of developing an SQT, which assures that the tests are fair, feasible and acceptable. Because of the strategic importance of Skill Qualification Tests to both training and personnel management, high level policy decisions were made about test content, validation, and scoring. The general requirements of the program are that tests must be (a) fair and feasible and (b) have validity demonstrated in advance of operational use.

FAIRNESS AND FEASIBILITY OF THE TESTS

Fairness means that all soldiers have an equal opportunity to demonstrate their true level of job competence. Test content must be based on actual job requirements, and testing conditions must be sufficiently constant throughout the Army so that scores obtained from administrations under varied conditions are not noticeably different. Tests given in

Alaska, Panama, and Korea must all be administered under similar conditions, and, in addition, all persons administering and scoring the tests must be able to do so accurately and objectively. An additional requirement is that the tests must be acceptable to soldiers and knowledgeable experts as fair measures of ability to perform critical job tasks. Therefore, fairness attends to requirements of both training and personnel management.

Feasibility requires that the tests be suitable for administration in all types of units; equipment, terrain, personnel, and all testing material must be readily available. Another aspect of feasibility is that testing time must be reasonable, with up to one day allowed for testing each soldier.

The requirements that Skill Qualification Tests be fair and feasible put severe limitations on the use of hands-on performance tests. The history of performance testing is that scoring accuracy and standardization are difficult to obtain. The resolution of the fairness and feasibility requirements is to have several kinds of testing. Under present policy decisions, all Skill Qualification Tests contain a written component, and some Skill Qualification Tests contain a hands-on component. Four hours of testing is allowed for the written component, and up to four hours is allowed for the hands-on portion. A third component, called performance certification, can also be included in Skill Qualification Tests. It is essentially an observational evaluation of actual job performance.

Therefore, an SQT may include up to three distinct types of tests, each with its own inherent strengths and weaknesses. A combination of these tests is the operational answer to the fairness and feasibility requirement.

Types of Tests. Hands-on performance tests are most desirable. They are a form of structured observation where a scorer evaluates an individual on a set of performance measures (observable behaviors). Advantages of hands-on testing are obvious: it tests actual performance, has high fidelity to the job, allows for immediate feedback, and has high face validity to examinees. However, considerable developmental effort is required to insure scoring reliability and standardization of conditions. It also is expensive in terms of equipment, personnel, and time, i.e., feasibility is often a problem. In order to ensure feasibility there is a natural tendency to truncate tests of tasks by shrinking the boundaries. Unfortunately, this may be at the expense of the validity of the test. For these reasons it is extremely difficult, if not impractical, to initiate a large-scale hands-on testing system for an organization as large as the Army. Therefore, a hands-on component constitutes a subset of an SQT.

An alternative form of hands-on testing is performance certification. The performance certification component covers tasks that are too long and/or complex to include in the hands-on component, and that do not lend themselves to testing in a written mode. Performance certification tests are to be administered and scored by soldiers' supervisors in the

normal job setting. Performance certification allows greater flexibility and avoids some of the feasibility problems encountered in a hands-on component. The greatest problems in performance certification are insuring reasonable standardization of job testing conditions across individuals and standardization of scoring by supervisors. Until sound methods are developed for addressing these problems, performance certification will remain a small portion of an SQT.

The decision to include a written component imposes careful consideration and analysis of what criterion-referenced measurement means in this context. Since the focus of Skill Qualification Tests is on ability to perform critical job tasks, that aspect must be retained. Each written test of a task is to consist of a set of items, where each item is designed to measure an essential behavior or step in performing the task. For tasks that require primarily mental skills, such as those in supply and administration, written tests of tasks are often similar to the behaviors required on the job, and the standards for ability to perform the test of the task can be reasonably close to those on the job. For other tasks that require psychomotor skills, written test items only simulate actual job behaviors, and the setting of realistic standards indicating ability to perform the task is a more arbitrary process. To help approximate realistic job conditions, written items may have multiple correct responses and a variable number of alternatives. This added flexibility increases the difficulty in developing appropriate methods for setting standards. The determination of reasonable standards for written tests of tasks is one of the most difficult issues in the SQT program.

Scoring the Tests. Because Army jobs and training programs are structured in terms of critical tasks, the appropriate level of scoring for the SQT should also be based on tasks. The concept of "scorable unit" was invented to help assure criterion-referenced measurement of task performance. A scorable unit is designed to measure ability to perform a specific task, or in the case of complex tasks, a well defined subtask.

Each written scorable unit consists of a set of items, where each item is designed to measure an essential behavior or step in performing the task. Each item is scored pass-fail, and a prescribed number of items must be passed to be GO on the written scorable unit. A GO is counted as ability to perform the task. The current resolution to setting standards for written scorable units is to require that an a priori number of items be passed. For example, if a scorable unit contains five items, then four must be passed to obtain a GO.

Hands-on and performance certification scorable units consist of a set of performance measures, where each performance measure is scored pass-fail, and a prescribed number of performance measures must be passed to be GO on the scorable unit. A GO on the scorable unit is interpreted as ability to perform the task. The standards of GO generally are comparable to what is required on the job.

The requirement that all scorable units be acceptable as fair measures of ability to perform tasks is applied to both the hands-on and written tests. Juries of experts must agree that the written items and hands-on performance measures reflect ability to perform the tasks. Perhaps a safer statement would be that failure to pass the items indicates that the person is not able to perform the task.

The most critical requirement of SQTs is their job relevance. The procedures for establishing job relevance are described in the following section.

ESTABLISHING A CORRESPONDENCE BETWEEN TEST CONTENT AND JOB TASKS

Test content of all SQTs is a sample of critical tasks from the domain of job tasks in the specialty. In this way the tests have a specifiable and explicit link to the job. For each Army job there exists a Soldier's Manual that lists the tasks for which a soldier in that specialty is responsible. Therefore, this set of tasks becomes the operational definition of the job. Tests to measure performance on specific job tasks listed in the Soldier's Manual are developed from appropriate task analyses, and the tests for each task are operational definitions of performance on the tasks. Performance on the individual tasks is summed to obtain a total score, which in turn serves as the operational definition of job competence. Modern instructional technology, with its emphasis on specification of objectives and verfication that those objectives are attained, supports the above process for establishing the content and focus of SQTs, and thereby lends added credibility to these procedures.

Though the task is the basic level of analysis, the validity of task proficiency measurement depends on the adequacy of the test of the task. By means of detailed task analyses, the set of performance measures or behaviors required for successful performance of the task are identified. These lists of performance measures are all available in the Soldier's Manual. Each item developed to test for task proficiency must occupy a clearly specified relationship to a performance measure required in task performance. Assuming that the set of items developed for a test of a task has been selected in accordance with the procedures described above, one may assume with reasonably high confidence that successful performance of each tested behavior is a necessary condition for successful performance of the task. How to score the set of items in a written scorable unit to obtain estimates of ability to perform tasks is a complex question. Measurement error is always a problem that must be allowed for. Whether being scored GO on a test of a task requires passing all items included in the test or something less than perfection depends on the nature of the task, the fidelity with which the task can be tested in a written mode, the complexity of the format (e.g., multiple correct responses), and the number of items within the cluster. Use of subject matter experts in reaching such a determination is mandatory.

In the case of a hands-on test of a task, measurement error arising from the use of words is minimized. However, other measurement problems arise. One is that a full performance test of a task generally is not feasible. It may be too costly in terms of time, equipment, or personnel. Therefore, a truncated test of the task is often developed by eliminating some of the performance measures or steps required for the full performance test. By truncating the test, though, it is possible that the tested portion is necessary to successful task performance, but is not sufficient.

VALIDATING TESTS PRIOR TO ADMINISTRATION

A first question to be resolved was how to define validity. The starting point was the usual definition of validity, i.e., that the tests measure what they are intended to measure. In the case of Skill Qualification Tests, the intent is to measure ability to perform critical job tasks. The content of the tests, therefore, becomes the crucial factor in establishing validity. The content must be thoroughly reviewed by experts to ensure that the right behaviors and decisions are assembled in each scorable unit. The first requirement, then, is consistent agreement among experts that the content of the test is based on ability to perform critical job tasks. A second requirement is that the scorable units discriminate between performers (masters) and nonperformers (non-masters). A third requirement applies only to written scorable units. All items in a written scorable unit must be consistent estimators of mastery on the task covered by the entire scorable unit. Thus, the conceptualizing of validity focuses on consistency: consistency between the content of the test and the job tasks, consistency among expert reviews, and consistency in identifying mastery.

Skill Qualification Tests are constructed and validated by Army agencies that have resident expertise in the job specialties. Generally these are the Army schools, but they also include other agencies, such as the Health Services Command. Since the test content must reflect job tasks, the test developers must have detailed task analyses available that identify the behaviors essential to successful performance of the tasks. Skill Qualification Tests are developed in the following conceptual sequence:

1. Identify tasks for testing;

2. Identify behaviors or steps essential for performing each task;

3. Develop scorable units to cover essential behaviors of the task, and review scorable units for content validity;

4. Try out scorable units on soldiers to verify accuracy of measurement.

After each step in the process, the products are submitted to higher headquarters for review and approval. The content of the scorable units is fixed after step 3. Scorable units found to be unsatisfactory through tryout on soldiers can be revised, but the content cannot be changed. Test content is fixed through agreement among experts that the contents of the scorable units are indeed valid measures of ability to perform the tasks, and the tryout serves only to establish the measurement properties of the scorable units.

The tryout with soldiers is different for the hands-on and written components. For the hands-on tests, the primary concern is to establish that the performance measures can be scored accurately. Acceptable agreement among the scores is considered to be attained when 80 percent of all pairs of rater scores are the same for the performance measures in a scorable unit. If less than 80 percent agreement is obtained, then the performance measures are revised until an adequate level of scoring consistency is attained.

For written tests the tryout is concerned with establishing the effectiveness of scorable units in distinguishing between performers and nonperformers, and with assuring that all elements in a scorable unit are consistent in estimating ability to perform the task. This tryout helps assure that all items of a scorable unit contribute to measuring performance on the task.

A final evaluation of the written scorable units is conducted after operational administration of the tests. A representative sample of answer sheets is selected for analysis, and the difficulty of items and scorable units are obtained. Those with high difficulty are examined to determine if they are faulty. Faulty items and scorable units are deleted prior to final scoring. When all steps of the review and analysis procedure for the written scorable units are accomplished, their validity as fair measures of ability to perform job tasks is considered to be reasonably well established.

ASSUMPTIONS FOR USING SQT SCORES

The assumptions on which SQTs are scored can be clearly explicated, as can the operations that determine test content, scoring procedures and standards.

In this section, three sets of assumptions made in using SQT scores are considered. These are using SQTs to (a) help determine training requirements, (b) help select soldiers in a single specialty, and (c) help select soldiers in merged specialties.

HELP DETERMINE TRAINING REQUIREMENTS

The assumptions required for using SQTs to help determine training requirements are straightforward. They are simply: (a) tasks can be defined--task elements or behaviors can be specified, conditions given, and standards of adequate performance established; (b) tasks can be measured validly--performance on the task is measured by scorable units, which contain time or performance measures related to task elements, and the sum of the elements passed in a scorable unit indicates quality of performance on the task; (c) task elements are weighted equally--items or performance measures corresponding to task elements or behaviors are scored as pass-fail, or as one-zero.

These three assumptions serve to provide operational definitions of performance on the tasks measured in SQTs. Although task elements do not have to be weighted equally, research evidence indicates that differential weighting generally does not improve the quality of measurement. A common practice is to give an element greater weight by preparing several items or performance measures for it.

The assumptions needed to help determine training requirements pertain only to tasks taken one at a time. Since the current training philosophy is to train on discrete tasks, no assumption about the interrelationships among the tasks is required.

HELP SELECT SOLDIERS IN A SINGLE SPECIALTY

The case of using SQTs to help select soldiers in a single job specialty requires additional assumptions about the interrelationships among job tasks and scorable units that measure task performance. The same three assumptions about measuring task performance are required (tasks can be defined, tasks can be measured validly, and task elements are weighted equally).

In addition, three more assumptions are required: (a) scorable units are weighted equally--all are scored as GO/NO-GO or as one-zero; (b) test score is the number of scorable units performed correctly--the total score is obtained by adding up the number of scorable units passed, and (c) the percent of scorable units passed indicates level of job performance--the percent of scorable units passed corresponds to the proportion of job tasks a soldier can perform. Given these assumptions, SQTs define the criterion of job proficiency, and the percent of scorable units correct (called percent-correct) is a direct reflection of job proficiency. Standards of job proficiency can then be set in terms of percent-correct scores.

HELP SELECT SOLDIERS IN MERGED SPECIALTIES

In the case of merged specialties, an additional assumption is required about the relationships among the jobs or groups of soldiers. The first six assumptions made in the case of the single specialty result in criterion-referenced measurement for each of the jobs being merged. However, in order to maintain criterion-referenced standards for merged specialties, the assumption is required that the jobs being merged are equal--that is, equal levels of proficiency in the individual jobs are equal to each other in an absolute sense, or stated operationally, all scorable units from all the relevant SQTs are weighted equally. Thus a soldier qualified in specialty 45N, for example, is equal to the qualified soldier in 45P, regardless of the percentage of soldiers in each qualified group. An implication of this assumption that the jobs being merged are equal is that if one qualified group contained 5 percent of a first MOS population while a second qualified group contained 50 percent of a second MOS population, the merged qualified group would contain proportionally more soldiers from the second group.

In the above example, each MOS would be represented in the merged qualified group in accordance with the number of soldiers from each MOS who attained qualifying scores. One MOS may be proportionally over-represented, while the second MOS is minimally represented or possibly not represented at all. How to use and maintain performance standards for merging MOS is a policy decision, and not a technical question. However, the criterion-referenced properties of SQTs permit rational policy decisions.

An alternative assumption in the case of merged specialties is that the groups, and not the MOS, are equal--that is, equal percentile-rank scores indicate equal levels of job proficiency. The use of percentile-rank scores, which indicate relative standing in a group, facilitates proportional representation of each MOS in the merged qualified group. For exmple, a policy decision could be made that 40 percent of each MOS be considered eligible for promotion. Such a policy decision might be made if policy makers were not willing to assume that the jobs were equal, or that the SQTs were not equally valid criterion-referenced measures of all the merged MOS, or if the policy makers decided that the need for proportional representation of the MOS in the qualified group outweighed the need to maintain performance standards. However, if SQTs are scored as percentile-rank and qualifications are based on percentile-rank scores, then the job performance standards would be given little or no consideration in determining the qualified group.

BENEFITS FROM USING CRITERION-REFERENCED SQTs

The change in focus from norm-referenced Military Occupational Specialty proficiency tests to criterion-referenced Skill Qualification Tests has enabled training and personnel management to obtain more comprehensive and meaningful information than before. Two major benefits that have resulted from the adoption of the criterion-referenced approach deal with (a) public nature of test content, and (b) job performance standards vs. personnel quotas. These benefits are discussed separately in the following paragraphs.

PUBLIC NATURE OF TEST CONTENT

An effective job proficiency testing program should be part of a larger system that includes job requirements and individual training programs. Modern instructional technology emphasizes the systems approach to training, and a job proficiency testing program is an integral component of the Army's modern training system.

Job requirements are defined by Soldier's Manuals, which list all the tasks a soldier in an MOS skill level (job) is responsible for performing. Soldier's Manuals are distributed throughout the Army for use by individual soldiers and for developing training programs, both resident courses and decentralized training conducted in units. Soldier's Manuals are also used to develop SQTs. No task can be tested that is not in the Soldier's Manual. Once the system becomes fully operational, all components of the Army can know what each soldier should be able to do, is able to do, and should be trained to do. There will be no surprise requirements.

In addition to Soldier's Manuals, soldiers are given additional detailed information about the job tasks on which they will be tested. This information is contained in the SQT Notice, which lists the specific tasks included in an SQT, how the tasks will be tested (written or hands-on), standards, and a description of the actual test content. Soldiers are given advance notice of what they will be required to know and do. All soldiers in an MOS are given equal information about what they will be tested on, potentially allowing them equal opportunity to prepare for the test. Test content, at least in general terms, is public knowledge.

The public nature of test content reduces the need for representative sampling of tasks. One reason representative sampling of tasks is important in the typical testing program is to give all examinees an equal opportunity to demonstrate their competence. With the SQT Notice, test content can be focused in special areas, such as areas that have high training needs or that are related to new equipment in the field.

The public nature of SQT content also helps establish an integrated training and testing program based on critical job requirements. By selecting test content that focuses on critical job requirements, training efforts will tend to be directed toward these same requirements. Thus, an integrated training and testing system is being developed based on job requirements.

As long as individuals are tested on the specific requirements of their jobs, there is no advantage to keeping the test content secret. In fact, if the test is directly related to performance on the job, then the proficient individual should already know the test content without the benefit of the information contained in a test notice.

Minimizing Effects of Job Assignments on Test Scores. A problem that arises in the typical testing program, where test content is kept secret, is that some individuals have special advantages over others. One possible advantage is that because of favorable job assignments, job tasks and test content are very closely related for some individuals. In the past soldiers who were working outside of their MOS were at a distinct disadvantage on the test content based on MOS-specific job tasks. The effects of bad assignments are minimized in the SQT program because all MOS soldiers are told specifically what content will be included in the test. The prior knowledge about test content tends to equalize opportunities.

In the past some soldiers have had advantages because they were more familiar with the voluminous references given for MOS tests. Some soldiers did not have the references available to them, and some even if they did, had difficulty in identifying the critical information within the mass of paper and words. In the Soldier's Manual and SQT Notices the critical information is distilled and made available to all MOS soldiers. Thus, soldiers with high verbal fluency or with access to specialized information no longer retain such a distinct advantage. Since the critical information is made available to all soldiers in a form readily understood, the opportunities to acquire competence are equally available to all soldiers.

Minimizing Fears About Taking Tests. Some individuals seem to have a knack for doing well on tests, while others seem to freeze when confronted with a testing situation. Test wiseness is frequently cited as an explanation of why some do better than expected, and test anxiety is ascribed as a reason why some do more poorly than expected. Both of these factors --test wiseness and test anxiety--are undesirable influences because they distort the meaning of test scores. In the SQT program where everyone has an opportunity to practice for the test, the effects of test wiseness and test anxiety are minimized, and the scores are more likely to reflect true levels of competence.

A factor related to test wiseness and test anxiety is the threat that many soldiers experience when taking tests. The threat may be viewed as having both objective and subjective components. A major source of objective threat arises from the fact that SQTs are used to help make personnel decisions that affect careers. Soldiers who do poorly on SQTs are likely to be penalized, while those who do well are rewarded. The test then, understandably, poses a threat to many soldiers, especially those who are marginal performers or who are not familiar with testing, or who have had negatively conditioning experiences in school situations. Subjective components of threat may arise from a variety of circumstances, such as personal characteristics, prior experience with tests, or from a fear of being evaluated. The fear of being evaluated may arise because the rules or basis for the evaluation are not explicit. If soldiers have foreknowledge about the tasks they will be evaluated on, and the means by which the evaluation will be conducted, then the subjective threat may often be reduced. Prior knowledge about test content may equalize opportunities for soldiers to demonstrate their true level of job competence by reducing distortion of test scores arising from subjective threat.

The public nature also has the general effect of increasing the validity of the tests. By giving all MOS soldiers more of an equal opportunity to prepare for the test, the test scores are more likely to reflect true levels of competence.

JOB PERFORMANCE STANDARDS VS. PERSONNEL QUOTAS

A criterion-referenced job proficiency test consisting of task-based tests can be scored in terms of percent of tests correct, which is a direct indicator of the percentage of job tasks a soldier can perform, and therefore, is a direct measure of level of job competence. The percent of task-based tests correct can be interpreted because standards are specified. The distribution of scores is not a relevant consideration in interpreting the meaning of the scores.[1]

---

[1] Norm-referenced proficiency tests, in which items have no meaning in terms of job-related activities, have meaning only in terms of percentile-rank scores. The percentage of items correct does not convey information because the population of items has not been defined precisely. Since such test scores have no external referent, the scores can be interpreted only in relation to the group taking that particular set of items. The tendency, based on traditional psychometric theory, is to select items on the basis of their difficulty and correlation with total test score. If items do not have the desired statistical properties, they are deleted or revised until they exhibit the proper difficulties and correlations with total score. Resulting changes in test content, and therefore, the correspondence between test and job content, are not systematically taken into account.

For each task in an SQT two categories of performance are established--qualified and not qualified. Therefore, SQTs provide GO/NO-GO decisions on task performance. Soldiers either meet these standards or they do not. The total SQT score is the sum of all scorable units passed, which provides continuous scores ranging from all scorable units correct to none, or 100 percent correct to 0 percent correct.

Current Army policy is that the SQT total score scale is divided into three categories. The higher passing score, called the Qualification Score, determines eligibility for award of the next higher skill level, and therefore eligibility for promotion. Only persons with the appropriate skill level are eligible for promotion. The Qualification Score is set at 80 percent of the scorable units correct. The lower passing score, called the Verification Score, determines eligibility to retain the current skill level; the Verification Score is set at 60 percent of the scorable units correct. Soldiers with SQT scores below 60 percent correct may be reclassified to another MOS.

Rank Ordering and Performance Categories. If SQT scores are also used to rank order soldiers, then in most cases the criterion-referenced power of the tests will be reduced or lost entirely. The following cases illustrate this point; the number of eligibles is a) equal to, b) less than, and c) greater than the quotas.

a) If the quotas and number of eligible soldiers are the same, then the decisions of whether to promote, based on the hurdle, and when to promote, based on rank order, have the same boundaries and there is no conflict between quotas and standards.

b) If the number of eligibles is less than the quota and the standards are waived until the quotas are met, then the rank ordering would be used to decide both whether and when to promote. Waiving standards could be equivalent to rank ordering. If the standards are waived one unit at a time until the quotas are satisfied, then the effect is to rank order with no regard to prerequisites. The waiving could be done in larger units, say from 80 correct to 60 correct, and then making the decision of when to promote on the basis of other factors. How the waiving is accomplished and how the tradeoff between standards and quotas is achieved, are policy decisions. Waiving standards forces an explicit decision about the tradeoff, whereas the pure rank ordering approach ignores any consideration of standards. On the other hand, if standards are not waived, then the rank ordering would be used only to decide when to promote. In this case the quotas would be waived in favor of increased quality.

c) If the number of eligibles is greater than the quota, then depending on how the pool of eligibles becomes replenished, the prerequisite standards may have varied meaning. If the pool of eligibles is always larger than the quota, then some soldiers near the cutting score may not be reached and consequently not promoted. If the pool is exhausted before new soldiers are added, then these soldiers are assured eventual promotion, and new soldiers who become eligible are placed into

a hold category until the original pool is exhausted. If the new eligible soldiers are immediately added to the pool, then there is no assurance that the remaining eligible soldiers from the original pool will be promoted even though they surpassed the prerequisite standards.

The main point about hurdles vs. rank ordering is that the criterion-referenced standards may be lost to the rank order unless explicit decisions are made to retain the standards. Rank ordering lends itself so easily to satisfying quotas that performance standards may be readily bypassed. The ability to obtain objective standards of job performance has profound impact on how personnel decisions can be made. Personnel managers have a choice between using a priori derived standards, independent of the population taking the test, and using quotas derived independent of the content of the test. The traditional solution to personnel decisions is to establish quotas, and then to select individuals until the quotas are satisfied.

According to the criterion-referenced test model, levels of performance within a proficiency category are not discriminated because the criterion levels are the only points of interest. Continuous scores are available, however, and they can be used for rank ordering soldiers. Because SQTs can be scored either in terms of performance categories or as continuous scores, explicit decisions can be made about which methods or combination of methods to use, and how the scores will be used in personnel decisions.

As a minimum, SQTs are used to set prerequisites for promotion. As described above, the prerequisite score is waived to meet quotas if such a policy decision is made. An immediate question is whether SQT scores should be used to rank order the pool of soldiers eligible for promotion. To oversimplify the question: SQTs are now used to determine whether to promote. The question of when to promote can also be answered on the basis of SQT scores, or can be based on other factors. (Other factors besides SQT scores do affect promotability, but the oversimplified version puts the issue in stark relief.) A discussion of how SQT scores can be combined with other factors is presented later in this section.

An unfortunate consequence of using quotas is that performance standards, which may be used in delineating a quota limit for one particular point in time, may not be entirely relevant when applied in another situation. If, for example, the top 50 percent in a job is eligible for promotion, the job performance of the eligible group will vary as the soldiers change over the years, or as the effectiveness of the training programs change, or as the relationship between test content and job requirements change over time.

Quality vs. Quantity in Personnel Decisions. A major breakthrough resulting from criterion-referenced SQTs is the availability of objective information about job competence that can be included in making personnel decisions. Level of job performance measured by these tests provides an absolute indication of proficiency that remains relatively constant as long as jobs remain defined by existing Soldier's Manuals. Performance

standards for personnel decisions can be specified in terms of the percentage of job tasks soldiers can perform. These standards are external to the test, and therefore more powerful statements can be made about the groups that are eligible to be selected in or out.

Quotas for personnel actions, such as promotion or attendance at a school, are likely to remain a driving force for personnel management in the foreseeable future. Rarely, if ever, will the number of soldiers eligible for a personnel action, based on performance standards, be the same as the required quota. Some adjustment to the quotas or performance standards, or both, generally will be required. If quotas are given top priority, then standards are waived; conversely, if performance is given top priority, then quotas are waived. If both quotas and performance are waived, say within some pre-established bounds, then a tradeoff between quality and quantity can be established.

Decision rules about quality vs. quantity can be explicitly stated. If performance standards are waived, there is a cost in terms of lowered individual performance (quality) in order to obtain sufficient numbers (quantity). If quotas are waived, there is a gain in individual performance (quality), but insufficient numbers (quantity) are obtained. By assigning values to units of performance and shortfalls, the tradeoff between quantity and quality can be calculated. Again, the tests do not dictate policy about quantity or quality, but they support decision rules and permit operations not possible without them.

Weighting Factors in Personnel Decisions. The situation becomes more complex when one does not base personnel decisions exclusively on test scores, but rather uses test scores as one factor in a composite score. Army personnel actions generally have been based on a composite score, which is characterized as the whole-man concept. The composites may be governed by explicit rules to provide objective indices, or the variables may be combined in a subjective manner by the decisionmakers. An example of explicit rules governing the combination of factors is Enlisted Evaluation Scores based on a weighting of MOS test scores and Enlisted Evaluation Report scores; another example is the determination of whether a soldier meets the prerequisites for a particular job training course, in which aptitude area scores, physical profile, and perhaps prior training may be considered. An example of subjective combination of factors is the process followed by a typical selection board that interviews soldiers, examines their records, and then arrives at a collective decision.

Criterion-referenced standards require the use of explicit rules for setting the minimum levels of qualification. If the process of combining scores for the qualified group is objective, explicit weights are assigned to each variable, and the contribution of each variable to the component score can be specified.

The assigned weights and the actual weights may or may not be the same. The actual weight of a factor is determined largely by the variability or range of scores for that factor. If the range is small, the effect is to add a virtual constant value to each individual's score, regardless of assigned weight, and the small differences can have only a small effect on the final rank ordering of the soldiers. If the combining is based on subjective judgment, then the weighting of the variables cannot be explicated. In either case, an important consideration is how the minimum qualifications are treated in determining eligibility for a personnel action. If the standards do serve to categorize soldiers into qualified and non-qualified groups and the qualified group is then given the favorable treatment while the non-qualified group is excluded from consideration, then the criterion-referenced standards are operative. If, however, the minimum standards can be waived, then the subjective process may easily ignore the standards, and the net effect may be to lose the power that inheres in criterion-referenced standards.

The process of combining scores may also be based on successive hurdles. The use of successive hurdles for combining scores virtually assures that standards will be maintained. Establishment of the minimum levels of qualifications requires explicit decisions, and any waiving then must also be explicit. An example of multiple hurdles is the determination of eligibility for entrance in a job training course. A minimum aptitude area score is set, usually at 90, and other minimum prerequisites may also be included in the decision, such as physical profiles, prior military job training, and high school courses completed. Not all eligible persons enter a course, but unqualified persons are excluded unless a specific waiver is applied. The use of hurdles is compatible with criterion-referenced standards.

SQTs, because of their criterion-referenced properties, permit basing personnel decisions on objective performance standards. As has been mentioned, technical feasibility does not necessarily dictate policy, and therefore personnel decisions need not necessarily be based on performance standards. However, since the possibility exists, rational evaluation of the costs and benefits in changing to new personnel policies can now be accomplished by decisionmakers.

CONCLUSIONS

Two themes have pervaded the discussion of criterion-referenced Skill Qualification Tests: 1) test content is based on systematic analysis of job requirements; 2) SQTs provide new opportunities for training managers, personnel managers, and research personnel to reassess and redefine their functions.

SQTs provide new information about levels of job performance not previously available from traditional proficiency and achievement tests. However, the power inherent in this information would be lost unless explicit use is made of the criterion-referenced performance data available from SQT scores.

For training managers and job supervisors, feedback from SQTs can be used to structure individualized training programs based on critical job tasks. Instead of basing training requirements on global evaluations of performance, training programs can be based on specific job tasks that are critical to both unit mission and individual job requirements.

Personnel managers have responsibilities for defining job specialties and for matching individuals and jobs. Under traditional procedures, jobs have tended to be defined in general terms of functions, skills, and knowledges. Similarly, individual qualifications have also been assessed in global terms, such as total MOS proficiency score, training courses completed, or time in grade. With the technology underlying the SQT program, and all of modern instructional technology, both job requirements and individual qualifications can be stated more precisely --critical job tasks define job requirements, and performance on these critical tasks defines levels of proficiency.

Finally, research personnel may have to reconceptualize their function. Traditionally, test psychologists have focused their efforts on developing statistical techniques for improving the accuracy of test scores. However, in criterion-referenced testing, establishing the content of a test is prerequisite to, and therefore, perhaps even more important than improving the accuracy of test scores. The interpretation of test scores in criterion-referenced testing is always dependent on being able to provide an explicit linkage between test content and test scores. Research efforts are required that explore and define the relationship between test content and test scores. For example, there is a need for research on development of score scales designed to reflect realistic standards of performance.

Because of the need to establish an operational testing program to meet a tight schedule, some decisions were made that appear reasonable but are not supported by an existing test theory. One example of such a decision is how to match scores from different tests. SQTs are assumed to be of equal difficulty and relevance to all job incumbents, which is a most reasonable assumption given the current state of the art. New theoretical developments are required to develop score scales that can equate scores of soldiers tested on different tasks. A promising approach is available in latent trait theory, which addresses many of the problems faced in developing SQTs. The applicability of latent trait theory, however, has not yet been sufficiently demonstrated in any large-scale testing program, especially one confronted with the limited resources available to test development activities.

DISTRIBUTION

ARI Distribution List

4 OASD (M&RA)
2 HQDA (DAMI-CSZ)
1 HQDA (DAPE-PBR
1 HQDA (DAMA-AR)
1 HQDA (DAPE-HRE-PO)
1 HQDA (SGRD-ID)
1 HQDA (DAMI-DOT-C)
1 HQDA (DAPC-PMZ-A)
1 HQDA (DACH-PPZ-A)
1 HQDA (DAPE-HRE)
1 HQDA (DAPE-MPO-C)
1 HQDA (DAPE-DW)
1 HQDA (DAPE-HRL)
1 HQDA (DAPE-CPS)
1 HQDA (DAFD-MFA)
1 HQDA (DARD-ARS-P)
1 HQDA (DAPC-PAS-A)
1 HQDA (DUSA-OR)
1 HQDA (DAMO-ROR)
1 HQDA (DASG)
1 HQDA (DA10-PI)
1 Chief, Consult Div (DA-OTSG), Adelphi, MD
1 Mil Asst. Hum Res, ODDR&E, OAD (E&LS)
1 HQ USARAL, APO Seattle, ATTN: ARAGP-R
1 HQ First Army, ATTN: AFKA-OI-TI
2 HQ Fifth Army, Ft Sam Houston
1 Dir, Army Stf Studies Ofc, ATTN: OAVCSA (DSP)
1 Ofc Chief of Stf, Studies Ofc
1 DCSPER, ATTN: CPS/OCP
1 The Army Lib, Pentagon, ATTN: RSB Chief
1 The Army Lib, Pentagon, ATTN: ANRAL
1 Ofc, Asst Sect of the Army (R&D)
1 Tech Support Ofc, OJCS
1 USASA, Arlington, ATTN: IARD-T
1 USA Rsch Ofc, Durham, ATTN: Life Sciences Dir
2 USARIEM, Natick, ATTN: SGRD-UE-CA
1 USATTC, Ft Clayton, ATTN: STETC-MO-A
1 USAIMA, Ft Bragg, ATTN: ATSU-CTD-OM
1 USAIMA, Ft Bragg, ATTN: Marquat Lib
1 US WAC Ctr & Sch, Ft McClellan, ATTN: Lib
1 US WAC Ctr & Sch, Ft McClellan, ATTN: Tng Dir
1 USA Quartermaster Sch, Ft Lee, ATTN: ATSM-TE
1 Intelligence Material Dev Ofc, EWL, Ft Holabird
1 USA SE Signal Sch, Ft Gordon, ATTN: ATSO-EA
1 USA Chaplain Ctr & Sch, Ft Hamilton, ATTN: ATSC-TE-RD
1 USATSCH, Ft Eustis, ATTN: Educ Advisor
1 USA War College, Carlisle Barracks, ATTN: Lib
2 WRAIR, Neuropsychiatry Div
1 DLI, SDA, Monterey
1 USA Concept Anal Agcy, Bethesda, ATTN: MOCA-WGC
1 USA Concept Anal Agcy, Bethesda, ATTN: MOCA-MR
1 USA Concept Anal Agcy, Bethesda, ATTN: MOCA-JF
1 USA Artic Test Ctr, APO Seattle, ATTN: STEAC-MO-ASL
1 USA Artic Test Ctr, APO Seattle, ATTN: AMSTE-PL-TS
1 USA Armament Cmd, Redstone Arsenal, ATTN: ATSK-TEM
1 USA Armament Cmd, Rock Island, ATTN: AMSAR-TDC
1 FAA-NAFEC, Atlantic City, ATTN: Library
1 FAA-NAFEC, Atlantic City, ATTN: Hum Engr Br
1 FAA Aeronautical Ctr, Oklahoma City, ATTN: AAC-44D
2 USA Fld Arty Sch, Ft Sill, ATTN: Library
1 USA Armor Sch, Ft Knox, ATTN: Library
1 USA Armor Sch, Ft Knox, ATTN: ATSB-DI-E
1 USA Armor Sch, Ft Knox, ATTN: ATSB-DT-TP
1 USA Armor Sch, Ft Knox, ATTN: ATSB-CD-AD
2 HQUSACDEC, Ft Ord, ATTN: Library
1 HQUSACDEC, Ft Ord, ATTN: ATEC-EX-E-Hum Factors
2 USAEEC, Ft Benjamin Harrison, ATTN: Library
1 USAPACDC, Ft Benjamin Harrison, ATTN: ATCP-HR
1 USA Comm-Elect Sch, Ft Monmouth, ATTN: ATSN-EA
1 USAEC, Ft Monmouth, ATTN: AMSEL-CT-HDP
1 USAEC, Ft Monmouth, ATTN: AMSEL-PA-P
1 USAEC, Ft Monmouth, ATTN: AMSEL-SI-CB
1 USAEC, Ft Monmouth, ATTN: C, Facl Dev Br
1 USA Materials Sys Anal Agcy, Aberdeen, ATTN: AMXSY-P
1 Edgewood Arsenal, Aberdeen, ATTN: SAREA-BL-H
1 USA Ord Ctr & Sch, Aberdeen, ATTN: ATSL-TEM-C
2 USA Hum Engr Lab, Aberdeen, ATTN: Library/Dir
1 USA Combat Arms Tng Bd, Ft Benning, ATTN: Ad Supervisor
1 USA Infantry Hum Rsch Unit, Ft Benning, ATTN: Chief
1 USA Infantry Bd, Ft Benning, ATTN: STEBC-TE-T
1 USASMA, Ft Bliss, ATTN: ATSS-LRC
1 USA Air Def Sch, Ft Bliss, ATTN: ATSA-CTD-ME
1 USA Air Def Sch, Ft Bliss, ATTN: Tech Lib
1 USA Air Def Bd, Ft Bliss, ATTN: FILES
1 USA Air Def Bd, Ft Bliss, ATTN: STEBD-PO
1 USA Cmd & General Stf College, Ft Leavenworth, ATTN: Lib
1 USA Cmd & General Stf College, Ft Leavenworth, ATTN: ATSW-SE-L
1 USA Cmd & General Stf College, Ft Leavenworth, ATTN: Ed Advisor
1 USA Combined Arms Cmbt Dev Act, Ft Leavenworth, ATTN: DepCdr
1 USA Combined Arms Cmbt Dev Act, Ft Leavenworth, ATTN: CCS
1 USA Combined Arms Cmbt Dev Act, Ft Leavenworth, ATTN: ATCASA
1 USA Combined Arms Cmbt Dev Act, Ft Leavenworth, ATTN: ATCACO-E
1 USA Combined Arms Cmbt Dev Act, Ft Leavenworth, ATTN: ATCACC-CI
1 USAECOM, Night Vision Lab, Ft Belvoir, ATTN: AMSEL-NV-SD
3 USA Computer Sys Cmd, Ft Belvoir, ATTN: Tech Library
1 USAMERDC, Ft Belvoir, ATTN: STSFB-DQ
1 USA Eng Sch, Ft Belvoir, ATTN: Library
1 USA Topographic Lab, Ft Belvoir, ATTN: ETL-TD-S
1 USA Topographic Lab, Ft Belvoir, ATTN: STINFO Center
1 USA Topographic Lab, Ft Belvoir, ATTN: ETL-GSL
1 USA Intelligence Ctr & Sch, Ft Huachuca, ATTN: CTD-MS
1 USA Intelligence Ctr & Sch, Ft Huachuca, ATTN: ATS-CTD-MS
1 USA Intelligence Ctr & Sch, Ft Huachuca, ATTN: ATSI-TE
1 USA Intelligence Ctr & Sch, Ft Huachuca, ATTN: ATSI-TEX-GS
1 USA Intelligence Ctr & Sch, Ft Huachuca, ATTN: ATSI-CTS-OR
1 USA Intelligence Ctr & Sch, Ft Huachuca, ATTN: ATSI-CTD-DT
1 USA Intelligence Ctr & Sch, Ft Huachuca, ATTN: ATSI-CTD-CS
1 USA Intelligence Ctr & Sch, Ft Huachuca, ATTN: DAS/SRD
1 USA Intelligence Ctr & Sch, Ft Huachuca, ATTN: ATSI-TEM
1 USA Intelligence Ctr & Sch, Ft Huachuca, ATTN: Library
1 CDR, HQ Ft Huachuca, ATTN: Tech Ref Div
2 CDR, USA Electronic Prvg Grd, ATTN: STEEP-MT-S
1 CDR, Project MASSTER, ATTN: Tech Info Center
1 Hq MASSTER, USATRADOC, LNO
1 Research Institute, HQ MASSTER, Ft Hood
1 USA Recruiting Cmd, Ft Sherdian, ATTN: USARCPM-P
1 Senior Army Adv., USAFAGOD/TAC, Elgin AF Aux Fld No. 9
1 HQ USARPAC, DCSPER, APO SF 96558, ATTN: GPPE-SE
1 Stimson Lib, Academy of Health Sciences, Ft Sam Houston
1 Marine Corps Inst., ATTN: Dean-MCI
1 HQUSMC, Commandant, ATTN: Code MTMT 51
1 HQUSMC, Commandant, ATTN: Code MPI-20
2 USCG Academy, New London, ATTN: Admission
2 USCG Academy, New London, ATTN: Library
1 USCG Training Ctr, NY, ATTN: CO
1 USCG Training Ctr, NY, ATTN: Educ Svc Ofc
1 USCG, Psychol Res Br, DC, ATTN: GP 1/62
1 HQ Mid-Range Br, MC Det, Quantico, ATTN: P&S Div

1 US Marine Corps Liaision Ofc, AMC, Alexandria, ATTN: AMCGS-F
1 USATRADOC, Ft Monroe, ATTN: ATRO-ED
6 USATRADOC, Ft Monroe, ATTN: ATPR-AD
1 USATRADOC, Ft Monroe, ATTN: ATTS-EA
1 USA Forces Cmd, Ft McPherson, ATTN: Library
2 USA Aviation Test Bd, Ft Rucker, ATTN: STEBG-PO
1 USA Agcy for Aviation Safety, Ft Rucker, ATTN: Library
1 USA Agcy for Aviation Safety, Ft Rucker, ATTN: Educ Advisor
1 USA Aviation Sch, Ft Rucker, ATTN: PO Drawer O
1 HQUSA Aviation Sys Cmd, St Louis, ATTN: AMSAV-ZDR
2 USA Aviation Sys Test Act, Edwards AFB, ATTN: SAVTE-T
1 USA Air Def Sch, Ft Bliss, ATTN: ATSA TEM
1 USA Air Mobility Rsch & Dev Lab, Moffett Fld, ATTN: SAVDL-AS
1 USA Aviation Sch, Res Trng Mgt, Ft Rucker, ATTN: ATST-T-RTM
1 USA Aviation Sch, CO, Ft Rucker, ATTN: ATST-D-A
1 HQ, USAMC, Alexandria, ATTN: AMXCD-TL
1 HQ, USAMC, Alexandria, ATTN: CDR
1 US Military Academy, West Point, ATTN: Serials Unit
1 US Military Academy, West Point, ATTN: Ofc of Milt Ldrshp
1 US Military Academy, West Point, ATTN: MAOR
1 USA Standardization Gp, UK, FPO NY, ATTN: MASE-GC
1 Ofc of Naval Rsch, Arlington, ATTN: Code 452
3 Ofc of Naval Rsch, Arlington, ATTN: Code 458
1 Ofc of Naval Rsch, Arlington, ATTN: Code 450
1 Ofc of Naval Rsch, Arlington, ATTN: Code 441
1 Naval Aerospc Med Res Lab, Pensacola, ATTN: Acous Sch Div
1 Naval Aerospc Med Res Lab, Pensacola, ATTN: Code L51
1 Naval Aerospc Med Res Lab, Pensacola, ATTN: Code L5
1 Chief of NavPers, ATTN: Pers-OR
1 NAVAIRSTA, Norfolk, ATTN: Safety Ctr
1 Nav Oceanographic, DC, ATTN: Code 6251, Charts & Tech
1 Center of Naval Anal, ATTN: Doc Ctr
1 NavAirSysCom, ATTN: AIR-5313C
1 Nav BuMed, ATTN: 713
1 NavHelicopterSubSqua 2, FPO SF 96601
1 AFHRL (FT) William AFB
1 AFHRL (TT) Lowry AFB
1 AFHRL (AS) WPAFB, OH
2 AFHRL (DOJZ) Brooks AFB
1 AFHRL (DOJN) Lackland AFB
1 HQUSAF (INYSD)
1 HQUSAF (DPXXA)
1 AFVTG (RD) Randolph AFB
3 AMRL (HE) WPAFB, OH
2 AF Inst of Tech, WPAFB, OH, ATTN: ENE/SL
1 ATC (XPTD) Randolph AFB
1 USAF AeroMed Lib, Brooks AFB (SUL-4), ATTN: DOC SEC
1 AFOSR (NL), Arlington
1 AF Log Cmd, McClellan AFB, ATTN: ALC/DPCRB
1 Air Force Academy, CO, ATTN: Dept of Bel Scn
5 NavPers & Dev Ctr, San Diego
2 Navy Med Neuropsychiatric Rsch Unit, San Diego
1 Nav Electronic Lab, San Diego, ATTN: Res Lab
1 Nav TrngCen, San Diego, ATTN: Code 9000-Lib
1 NavPostGraSch, Monterey, ATTN: Code 55Aa
1 NavPostGraSch, Monterey, ATTN: Code 2124
1 NavTrngEquipCtr, Orlando, ATTN: Tech Lib
1 US Dept of Labor, DC, ATTN: Manpower Admin
1 US Dept of Justice, DC, ATTN: Drug Enforce Admin
1 Nat Bur of Standards, DC, ATTN: Computer Info Section
1 Nat Clearing House for MH Info, Rockville
1 Denver Federal Ctr, Lakewood, ATTN: BLM
12 Defense Documentation Center
4 Dir Psych, Army Hq, Russell Ofcs, Canberra
1 Scientific Advsr, Mil Bd, Army Hq, Russell Ofcs, Canberra
1 Mil and Air Attache, Austrian Embassy
1 Centre de Recherche Des Facteurs, Humaine de la Defense Nationale, Brussels
2 Canadian Joint Staff Washington
1 C/Air Staff, Royal Canadian AF, ATTN: Pers Std Anal Br
3 Chief, Canadian Def Rsch Staff, ATTN: C/CRDS(W)
4 British Def Staff, British Embassy, Washington
1 Def & Civil Inst of Enviro Medicine, Canada
1 AIR CRESS, Kensington, ATTN: Info Sys Br
1 Militaerpsykologisk Tjeneste, Copehagen
1 Military Attache, French Embassy, ATTN: Doc Sec
1 Medecin Chef, C.E.R.P.A.-Arsenal, Toulon/Naval France
1 Prin Scientific Off, Appl Hum Engr Rsch Div, Ministry of Defense, New Delhi
1 Pers Rsch Ofc Library, AKA, Israel Defense Forces
1 Ministeris van Defensie, DOOP/KL Afd Sociaal Psychologische Zaken, The Hague, Netherlands